

Submitted for publication in Economic Notes  
CREDIT RISK MODELS  
DO THEY DELIVER THEIR PROMISES?  
A QUANTITATIVE ASSESSMENT

November 13, 2002

GIANLUCA ODERDA <sup>1</sup>  
*Pictet & Cie*  
*29 Bd-Georges Favon*  
*CH 1204, Geneva, Switzerland*

MICHEL M. DACOROGNA <sup>2</sup>  
*Converium AG*  
*General Guisan Quai 26*  
*CH 8022 Zurich, Switzerland*

TOBIAS JUNG <sup>3</sup>  
*Zurich Insurance Company*  
*International Program Business Switzerland*  
*Austrasse 46,*  
*CH-8045 Zurich, Switzerland*

**Abstract**

We develop a framework to assess the statistical significance of expected default frequency calculated by credit risk models. This framework is then used to analyze the quality of two commercially available models that have become popular among practitioners: KMV Credit Monitor and RiskCalc from Moody's.

Using a unique database of expected default probability from both vendors, we study both the consistency of the prediction and its timeliness. We introduce the concept of cumulative accuracy profile (CAP) that allows to see in one curve the percentage of defaulting companies captured by the models one year in advance. We also use the Miller's information test to see if the models add information to the S&P rating.

The result of the analysis indicates that these models indeed add relevant information not accounted for by rating alone. Moreover, with respect to rating agencies, the models predict defaults more than ten months in advance on average.

---

<sup>1</sup>Email address: [goderda@pictet.com](mailto:goderda@pictet.com)

<sup>2</sup>Corresponding author, Email address: [michel.dacorogna@converium.com](mailto:michel.dacorogna@converium.com)

<sup>3</sup>Email address: [tobias.jung@zurich.com](mailto:tobias.jung@zurich.com)

# 1 Introduction

The three pillars of the recently reinforced Basel Capital Accord, which by 2004 will be adopted by regulators in most industrialized countries, are minimum capital requirements, supervisory review of an institution's capital adequacy and internal assessment process, and market discipline through disclosure of banking practices. In particular, with the Basel Capital Accord of 1998, banks around the world have been allowed to assess regulatory capital issues related to credit risk by means of internal models. At the same time the development of the securitization of bond portfolios has brought to light the need for quantitative estimation of credit risks. Both phenomena largely explain why credit risk modeling has been evolving at a rapid pace over the last few years, and many commercially available models have appeared on the market. Among the latest releases we find RiskCalc, a tool by Moody's Risk Management Services (RMS), which is based both on a structural Merton approach [1] and on a statistical non-linear multi-factor model, generalizing early ideas of Altman [2]. For more information about commercially available credit risk models, we refer the reader to the excellent review paper by Crouhy *et al.* [3].

The purpose of this paper is twofold. On the one hand, to propose statistical methods to quantitatively evaluate the quality of credit risk models. On the other hand, to use them for studying and assessing two commercially available systems: the KMV Credit Monitor (KMV) and Moody's Risk Calc<sup>4</sup>. We perform a rolling statistical analysis in time of the predicting power of both models. To our knowledge, such kind of analysis has never been attempted before by an independent group. The purpose is not to rate the models against each other but rather to assess their usefulness. Indeed, we shall show in this paper that both models contain important information for predicting the risk of a company defaulting. Thus, they can provide at the same time early warning signals and quantitative risk assessment to credit and surety underwriters.

The work we present here required serious efforts from our side both in terms of understanding the data provided by the two companies and in terms of processing them in a consistent way. The outline of the paper reflects the procedure we followed: we start by discussing the input default probability and credit rating data, their sources and the criteria we use to interpret them. Then we illustrate the methodology followed in the analysis, presenting in detail the nature of the tests performed and their results.

## 2 Structural vs. Hybrid Credit Risk Models

As a brief theoretical review, the idea underlying the Merton model [1] is that, from the shareholder's point of view, the equity value of a company,  $E$ , can be treated as a

---

<sup>4</sup>It is interesting to note that since we did this study in the summer of 2001 Moody's has bought KMV and the two quantitative groups are now integrated.

call option written on the value of assets,  $A$ , where the strike price,  $B$ , is given by all the short-term liabilities,  $B_S$ , plus half of the long-term ones,  $B_L$ . The isomorphism between the market value position of equity holders and the position of holders of a call option on the firm's assets allows to write down the set of equations

$$\begin{aligned} E &= f(A, \sigma_A, r, B, T) \\ \sigma_E &= g(\sigma_A), \end{aligned} \tag{1}$$

where  $\sigma_A$  and  $\sigma_E$  are asset and equity volatility respectively,  $r$  is the risk free interest rate, and  $T$  is the time horizon of the option holder. From the two above equations it is possible to obtain asset value and volatility, otherwise unobservable quantities on a regular, short-term basis, from equity value and volatility, which are easily available from stock market data. It provides a convenient framework to get rid of the risk premium contained in the share price. Once the asset value and its volatility are obtained, a distance to distress (or distance to default) can be determined as a multiple of asset volatility,

$$DD = \frac{A - (B_S + \frac{1}{2}B_L)}{\sigma_A}. \tag{2}$$

It is clear from this equation that a default situation is approached when the expected value of assets becomes closer and closer to the expected value of liabilities, and when the asset volatility is large, therefore making it possible, from a statistical point of view, to have insufficient assets backing liabilities. Before the Merton model was formulated, credit risk was dealt with by means of multi-factor analysis, following early work by Altman [2]. The Altman Z-score model is a classification of corporate borrowers, which can also be used to get a default probability prediction, based on statistics performed on a sample of companies. The Z-score is defined as a linear combination of accounting ratios. The linearity of the model has been criticized, because the path to bankruptcy can be highly non-linear. Furthermore, since the model is entirely based on accounting ratios, which appear only at discrete quarterly intervals, its ability to pick up the risk of defaults in rapidly deteriorating conditions has also been questioned. Because of the many open issues left on the table, the Altman model has certainly contributed to open the way to the theory of credit risk modelling and, as we will see below, it seems to inspire part of the ideas underlying Moody's RiskCalc.

One of the commercial model making full and consistent use of the Merton approach is KMV Credit Monitor [4]. Based on historical Moody's default studies on large samples of firms, the KMV model establishes on the one hand a functional relationship between default probability and distance to distress, while, on the other hand, it defines per each official credit rating group (the rating being assigned by an agency such as Standard & Poors - S&P - or Moody's) an average default probability. The overall result is a mapping between distance to distress and implied credit rating. The KMV approach has been proved to be able to predict default earlier than rating

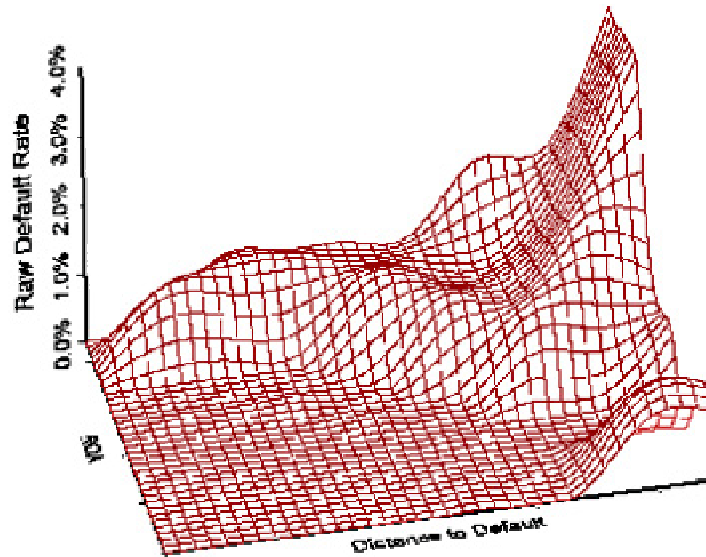


Figure 1: Correlation between the Return on Asset (RoA) and the Raw Default Rate (source Moody's).

agencies, as we will also show below. However, very recently, objections have been raised about purely structural models, and their ability at providing accurate measures of default risk has been questioned. Among others, CreditRisk+ [5] developed by Credit Suisse Financial Products proposes an actuarial approach instead of the usual structural model.

The new Moody's tool, RiskCalc, is based on a hybrid approach. It relies on the observation that, if we plot the relationship between distance to default and default probability for sets of companies having different levels of Return on Assets (ROA) we find the very irregular surface of Figure 1. If ROA were not at all a driver of default probability, we would find a much smoother result coming from the exact repetition along the ROA direction of the same curve defined in the Default Probability-Distance to Default plane.

Nine factors have been found to be relevantly correlated to default probability. They are:

1. Distance to Default, from the Merton model
2. Official Agency Rating
3. Return on Assets
4. Return on Equity
5. Size, defined by the amount of Total Assets
6. Liquidity, defined by the ratio of Working Capital over Total Assets
7. Leverage, defined by the ratio of Liabilities over Assets

<b>Universe</b>	<b>KMV Credit Monitor</b>	<b>Moody's RiskCalc</b>
North America	10021	10038
Europe	6658	4442
Financials	5564	
Asia	6831	

Table 1: Number of companies in the available datasets of KMV and Moody's

8. Equity Volatility

9. Equity Trend

By looking at the historical database of public companies, default probability is non-linearly regressed vs. the above set of factors. At the same time, to each official Moody's rating group a specific range (or value) of default probability is associated. As for the KMV model, a mapping is thus obtained between default probability and implied rating. A recent study by KMV [6] has put in doubt the relationship between ROA and default probability. Based on a statistical study on Moody's cohorts, KMV could not find any added value in having supplemental ROA information, compared to KMV expected default probability information alone. Their conclusion is that the implementation of the Merton model by Moody's differs from their own. Unfortunately, their comparative study is limited since they did not have access to the full results of Moody's model. In the following sections, we directly study the output of both KMV Credit Monitor and Moody's RiskCalc to see if they both add value and how they compare to each other.

## 3 Input Data

### 3.1 Processing Data Universes

Both KMV Credit Monitor and Moody's RiskCalc provide databases with default probability data. KMV Credit Monitor provides as well official S&P rating data, although for a very limited subset of companies. In both cases data are reported on a monthly basis for a total of 60 observations, ranging, for the purpose of this analysis, from April 1996 to March 2001.

Default probabilities give us a prediction with a 12-month validity in the future. Table 1 below gives an overview of the number of companies within the available datasets. For North America, 9039 companies - about 90% of the entire universe - could be identified in both the KMV and the Moody's databases by looking at a combination of Company Name and Stock Market Ticker. These companies constitute a common subset, on which the two models can be compared directly. Out of the 1935 S&P-rated North American Companies available in the KMV database, 1802 are also present in the Moody's database. It should be pointed out that the comparison of

KMV System Code	S&P Rating	KMV System Code	S&P Rating
0	No indication	13	BB+
1	Unused	14	BB
2	AAA	15	BB-
3	Unused	16	B+
4	AA+	17	B
5	AA	18	B-
6	AA-	19	CCC
7	A+	20	CC
8	A	21	C
9	A-	22	D
10	BBB+	23	CCC+
11	BBB	24	CCC-
12	BBB-	25	No indication

Table 2: Numerical encoding of S&P ratings in the KMV Database

the two models can only be performed on the North American database, since for the European universe the amount of S&P rated companies is too low to be statistically meaningful. We have stored default probabilities and, where available, associated S&P-ratings, into Microsoft Excel-files. In particular, we have downloaded Moody's Riskcalc data from the website. KMV Credit Monitor data are instead available directly from the program configuration. Before analyzing data quantitatively, we had to understand them. Since all the information, including the S&P rating, is numerically encoded, we proceed as follows:

1. We find the numerical equivalent of each S&P rating. Results are reported in Table 2 below.
2. As discussed in the introduction, both KMV and Moody's establish a correspondence between different levels of default probabilities and implied ratings, based on historical default information of large samples of firms. We are interested in the default probabilities which are mapped to an S&P D-rating and to a Moody's Caa rating (those are the ratings assigned to close-to-insolvent obligors). We find that such probability amounts to 20% for KMV, and 10.95%<sup>5</sup> or higher for Moody's (which is the cut-off point for the Caa class of Moody's implied rating. It corresponds to obligors close to default).

---

<sup>5</sup>the Caa cutoff point was obtained from private communication with Mr. B. Khandani (Moody's RMS)

## 3.2 Data cleaning

In order to obtain a consistent data set, several exclusions both for EDF-values and S&P ratings must be done:

1. Zero values of the default probability indicate that a company is not being followed on a certain month. This might be due either to a lack of coverage or to an exclusion following default. We analyse iteratively every single zero value inside the database and decide to replace it with the S&P D- or Moody's Caa-rating default probability (20% and 10.95% for KMV and Moody's, respectively) when the zero immediately follows an observation of D - or Caa - rating default probability.
2. Values of S&P rating equal to either 0 or 25 indicate once again a lack of coverage on a certain month. Following the above reasoning, we have decided that a 0 or a 25 following a 22 value of the S&P rating, which corresponds to D, has to be replaced by 22, since it reflects the situation of a defaulted obligor, which is not any more covered by the rating agency.

These anomalous values of either default probability or S&P rating, however, are not too numerous.

# 4 Data analysis: Methodology and Results

## 4.1 Model Predicted vs. Actual Defaults

We look at the predictive power of the two examined models by considering the subset of S&P rated companies, where the default event is identified by the S&P rating of D (numerically equivalent to 22, as in Table 2). We divide the sample into  $N = 6$  default probability intervals whose spacing grows exponentially <sup>6</sup>, as in Table 3 <sup>7</sup>. We count the number of companies  $C_{ki}$  falling into each interval,  $I_k$ ,  $k = 1, \dots, N$  at time  $T_i$ ,  $i = 1, \dots, 48$ . We limit ourselves to the 48<sup>th</sup> observation, because we need to compare each model prediction with the actual number of occurred defaults 12 months later, and we have a total of 60 available observations. For each interval, we measure the average default probability,  $\mu_{ki}$ . We define the product

$$\hat{D}_{ki} = \mu_{ki} C_{ki}, \quad (3)$$

as the expected number of defaults to be observed after 12 months, at time  $T_{i+12}$ . We exclude from the analysis at each time  $T_i$  obligors having a D S&P rating (for

---

<sup>6</sup>Exponentially spaced intervals ensure a statistically adequate and more homogeneous number of firms in each class, since the distribution of default probabilities is itself exponential.

<sup>7</sup>Courtesy of Moody's RMS

Financial Security	Moody's Class	Default Probability Classes	
		Lower Bound	Upper Bound
Exceptional-Excellent	Aaa-Aa3	0.00%	0.07%
Good	A1-A3	0.07%	0.15%
Adequate	Baa1-Baa3	0.15%	0.73%
Questionable	Ba1-Ba3	0.73%	3.16%
Poor	B1-B3	3.16%	10.95%
Very Poor	Caa1 and worst	10.95%	25.00%

Table 3: Default probability classes used for the analysis.

these the default event is, in fact, certain). For the companies belonging to each class, we read off their S&P rating after 12 months and count the number of obligors  $D_{ki}$  having reached a D S&P rating (numerically equivalent to 22). We then test the hypothesis

$$\hat{D}_{ki} = D_{ki}, \quad (4)$$

by building the  $\chi^2$ -distributed variable [8]

$$\chi^2 = \sum_{k=1}^N \frac{(\hat{D}_{ki} - D_{ki})^2}{(\hat{D}_{ki} + D_{ki})}. \quad (5)$$

This test can be used for the purpose of comparing the two models, by looking at which has the lowest  $\chi^2$  per degree of freedom. In our case the number of degrees of freedom (DF) is equivalent to the number of classes, 6 [8], and sets the critical  $\chi^2$  value to  $\chi_c^2 = 12.52$ .

The results we have found are reported in Appendix A respectively in Table 4 for the universe of S&P-rated companies in the KMV North American database, and in Table 5 for the universe of S&P-rated companies in the Moody's North American database. We have deliberately omitted the first 21 months of history, since too few default events were observed during this time. The  $\chi^2$  values per observation are plotted in Fig. 2 for each of the two models, together with the critical value  $\chi_c^2$ , represented by the constant dashed line. We see that, although both models give acceptable results from a statistical perspective, KMV Credit Monitor seems to be superior, particularly for the last few observations, referring to the time window between the last quarter 2000 and the first quarter 2001, when the North American economy started to deteriorate and, correspondingly, the number of defaults to increase. Notice also, from Tables 4 and 5 in Appendix A that the KMV Credit Monitor predictions are superior to the Moody's RiskCalc ones especially in the last two classes, corresponding to the most dangerous ranges of default probabilities (see Table 3).

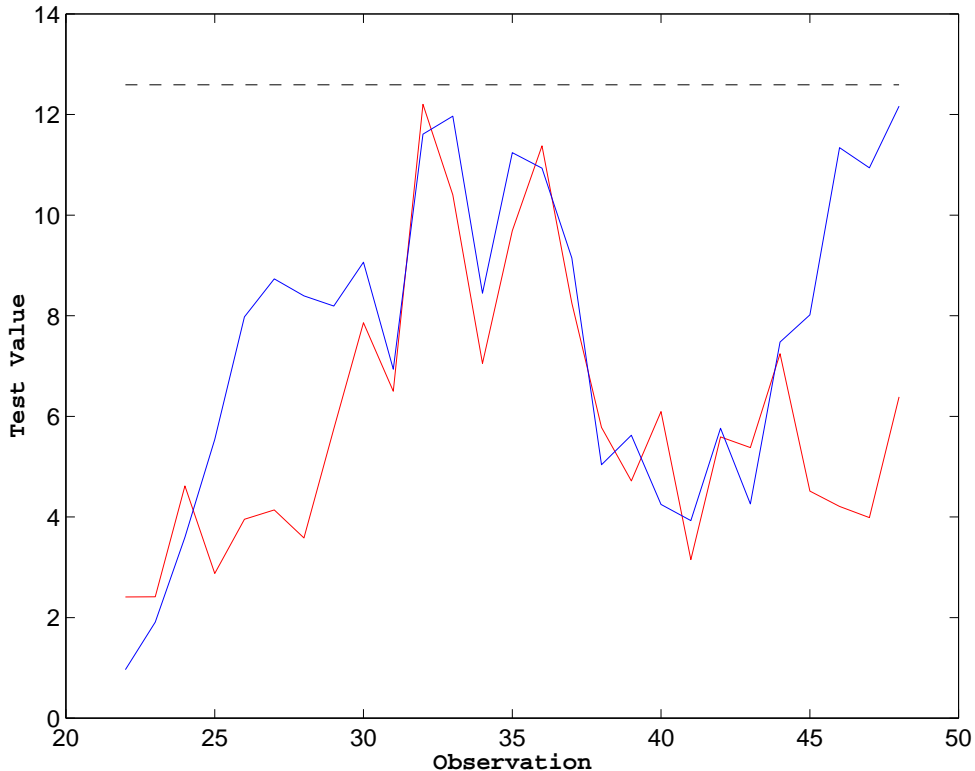


Figure 2:  $\chi^2$  test results per observation for KMV (red line) and Moody's (blue line). The dashed line represents the critical value,  $\chi_c^2 = 12.52$  for the 6 DF available.

## 4.2 CAP Profiles

For all the available observations, we build Cumulative Accuracy Profiles (CAPs). To do so, at each time  $T_i$ ,  $i = 1, \dots, 48$ , we order companies by default probabilities, from riskiest to safest. Then we consider the  $X\%$  riskiest companies, varying  $X$  between 0% and 100%, and check, 12 months later, what fraction  $Y(X)$  of defaulters (default is defined again from the explicit S&P rating of D) comes from this subset. For a model assigning random default probabilities, we should approximately find the line  $Y(X) = X$ . The more informative the model, the quicker the quantity should approach 100%, and therefore the steeper the CAP curve. For a perfect model, the CAP curve looks like a line capturing 100% of the defaulters with fraction of the population exactly equal to the default rate of the sample. We take the mean over all the observations, starting from observation 22, in order to have enough statistics for the CAP Profiles to make sense (we recall from the analysis in Section 4.1 that for the first 21 observations a very limited number of defaults was reported).

Again, the results are not too different, although it can be clearly noticed that the KMV CAP Profile is steeper than the Moody's one, especially in the percentiles corresponding to the riskiest companies. In other words, the KMV model seems to be slightly better at pinning down risky credits. This confirms what we pointed out

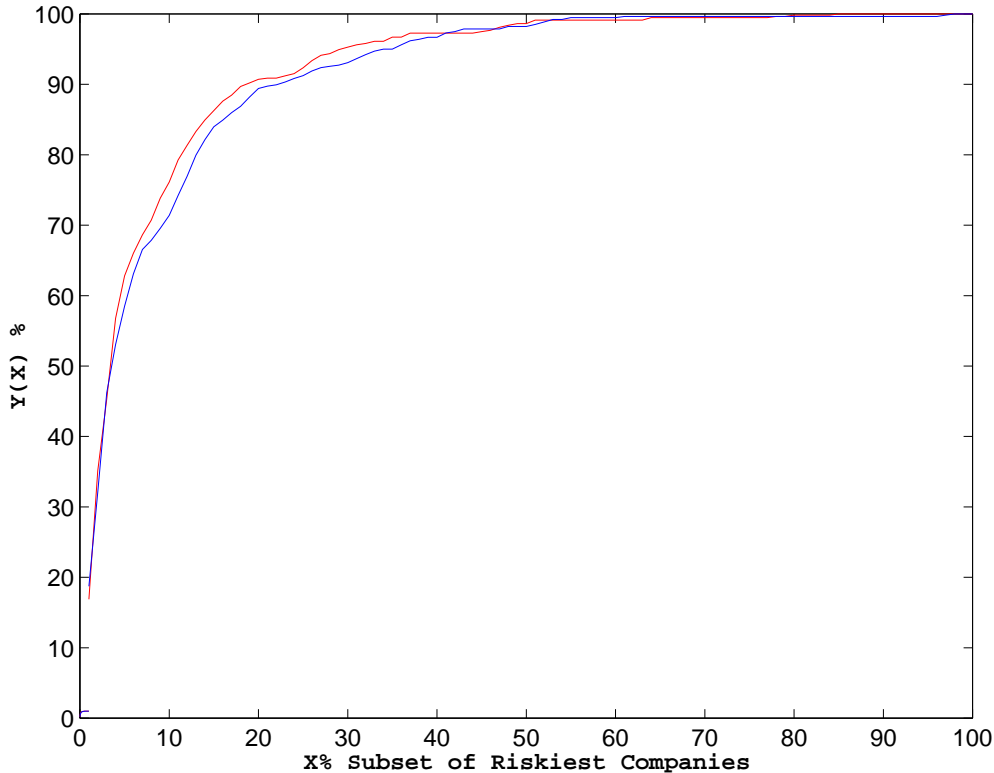


Figure 3: CAP profiles calculated for KMV (red line), and for Moody's (blue line).

at the end of Section 4.1, where the KMV superiority in the most dangerous default probability classes was already identified.

We can associate to each CAP profile an accuracy ratio (AR), defined as follows:

$$AR = \frac{2 \int_0^1 Y(X)dX - 1}{1 - f}, \quad (6)$$

where  $Y(X)$  represents the CAP profile function, and  $f$  is the actual (observed) fraction of defaults. Graphically the numerator is twice the area delimited from above by the CAP profile and from below by the line  $Y(X) = X$ , which is - we recall from above - the CAP profile of a model with no predictive value. The denominator, on the other hand, is twice the area delimited from below by the same boundary, and from above by the CAP profile of the ideal model, which would correctly identify the first  $f\%$  of the universe as the subset originating exactly 100% of the defaults (and therefore could be graphically visualized by a step-like function, connecting the origin to the point with coordinates  $(f\%, 100\%)$ , and then remaining flat at 100%). For the two models, the results we have found are

$$\begin{aligned} \text{AR}^{\text{KMV}} &= 87\% \\ \text{AR}^{\text{Moody's}} &= 85\%, \end{aligned} \tag{7}$$

and seem to be in line with what reported by Moody's in Ref. citeMoody'sRat-Methodology, in the range between 70% and 80%.

### 4.3 Comparison of Model Promptness at Signalling Defaults

We work with the subset of S&P rated companies available both in the KMV and the Moody's databases. Starting from observation 24 onwards, we take all the S&P default-rated companies, record the time of default,  $T_D^{\text{S\&P}}$ , and the time at which the threshold value of default probability, implying a distressed financial situation, is reached, respectively  $T_D^{\text{KMV}}$  and  $T_D^{\text{Moody's}}$ . We choose to identify the credit distress threshold by the cutoff value of 10.95% in Table 3, corresponding to the worst class of Moody's ratings. This way, we can quantitatively measure how prompt are the two models at detecting bad obligors. We measure

$$\begin{aligned} \Delta^{\text{KMV}} &= T_D^{\text{S\&P}} - T_D^{\text{KMV}} \\ \Delta^{\text{Moody's}} &= T_D^{\text{S\&P}} - T_D^{\text{Moody's}}. \end{aligned} \tag{8}$$

We find that a total of 109 companies, out of the 1802 in the common database, receive an S&P D-rating at some point during the 60 months of observations. For each of these companies, we evaluate the quantities in Eq. (8). The results are reported in Table 6 in Appendix A. In both cases we find a positive lag, meaning that the two models lead S&P rating at signalling defaults. The average and median (expressed in months) for the KMV lag are:

$$\begin{aligned} \mu^{\text{KMV}} &= 15.15 \\ m^{\text{KMV}} &= 12, \end{aligned} \tag{9}$$

whereas for Moody's RiskCalc we find

$$\begin{aligned} \mu^{\text{Moody's}} &= 10.91 \\ m^{\text{Moody's}} &= 7. \end{aligned} \tag{10}$$

Both models therefore are capable to lead default events, although KMV seems to be superior to RiskCalc in the prediction timeliness.

Using the 1802 companies in the common database, we have also tried to determine the number of false default alarms we have from the models. For each company reaching the credit distress threshold, we have checked if a default event really happens within 12 months from the model warning signal. We have repeated the analysis for each of the observations available, measuring the ratio

$$\rho_i = \frac{D_{i,\text{Actual}}}{D_{i,\text{Model}}}. \quad (11)$$

This ratio expresses the percentage of justified default alarms provided by the models. Assuming that model calibration has been performed correctly, by looking at Table 3, we would expect the average ratio in Eq. 11 to be between 10.95% and 25%. For the two models the average values found are respectively:

$$\begin{aligned} \rho^{KMV} &= 12.74\% \\ \rho^{Moody's} &= 12.00\%, \end{aligned} \quad (12)$$

confirming our prediction. Once again, we see that the models give comparable answers. Both produce a false default alarm in about 88% of the cases.

#### 4.4 Miller Test

To further assess the prediction power of the two models, we perform the non parametric test proposed by Miller [7]. Let us consider two risk models,  $M_A$  and  $M_B$ . At each time  $T_i$ , we take the population of firms and sort them into groups according to the  $M_A$  measure of default risk, so that, in the same group, all firms would have similar  $M_A$ -defined default risk. We call these groups  $M_A$  cohorts. If the  $M_B$  measure of default risk has predictive information not contained in the  $M_A$  measure, we should be able to find it in the following fashion. Within each of the  $M_A$  cohorts, we sort by values of the  $M_B$  measure. If the latter has additional power, there will be a higher default rate for the low quality firms within each  $M_A$  cohort, as determined by the  $M_B$  measure of default risk. In other words, the probability of default will increase for increasing values of the  $M_B$  default risk measure. Conversely, if no additional power is found, the actual defaults should be randomly scattered within a cohort. The behavior can be identified by means of the Kolmogorov-Smirnov (KS) test. For each  $M_A$  cohort, a CAP profile is built which relies on  $M_B$  default probabilities. The KS test works by calculating the maximal distance between the CAP profile and the uniformly distributed CAP profile, which should occur if the  $M_B$  default risk measure did not add any value. We perform the Miller test to evaluate the effectiveness of the default indicators provided both by KMV Credit Monitor and by Moody's Risk Calc vs. the official agency S&P rating. To this purpose, we sort the population of S&P-rated obligors present in both the KMV and Moody's databases into cohorts, defined according to the respective default indicators, and study the CAP profiles built on S&P rating. Then we do the reverse, by sorting the population into S&P rating cohorts, and by studying the CAP profiles built respectively on the KMV and

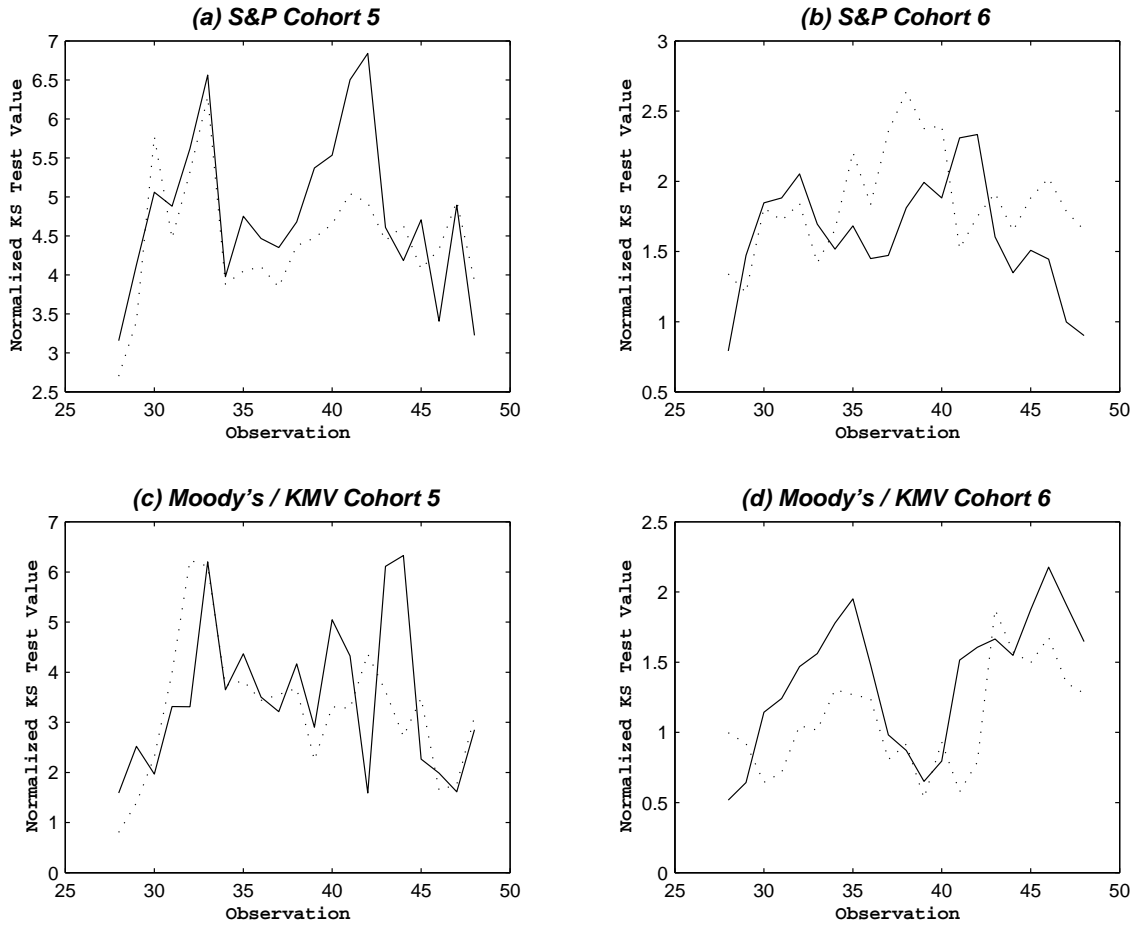


Figure 4: Miller test results. Plots (a) and (b) show the normalized KS test values per observation for Cohorts 5 and 6 respectively, defined according to S&P rating. The test is performed by taking as  $M_B$  each of the credit models under study, whose information added value is measured. The dashed line corresponds to Moody's RiskCalc, the continuous line to KMVCredit Monitor. Plots (c) and (d) show the normalized KS test values per observation for Cohorts 5 and 6 respectively, defined according to KMV (continuous line) and Moody's (dashed line) default probability. The test is performed by taking as  $M_B$  the S&P agency rating, whose information added value is measured.

Moody's risk measures. In order to have sufficient statistics, we only look at the two worst cohorts, since they originate almost all of the default events. For each case, we evaluate the ratio between the calculated KS test and the critical value, taken with a 95% confidence level.

We plot the results we have found in Fig. 4. The value added by the two credit models with respect to the official S&P agency rating are presented in Figs. 4a and 4b. The dashed line refers to Moody's, the continuous line to KMV. We see that, for all the observations, the normalized KS test values are much higher than the critical boundary of 1, below which the models would add no value to the information already contained in the rating. It is also important to notice that the KS test values are

much higher for cohort 5 than for cohort 6 (compare Figs. 4a and 4b.). We interpret this as a positive indication of the value added by the credit models, which is higher for obligors not yet classified in the worst rating class than it is for obligors already marked as mostly dangerous by the rating agency.

Conversely, we present in Figs. 4c-4d the value added by S&P to the two credit models. They indicate that for most observations and for both the cohorts studied, rating seems to add more value to the KMV model than to the Moody's RiskCalc model, which is exactly what we would expect. In fact some of the financial ratios underlying the credit assessment by the rating agency are accounted for in the Moody's RiskCalc Credit Risk measure, as discussed in Section 2 above. This is particularly evident for Cohort 6, where, for most of the observations, the normalized KS test is below the significance boundary of 1.

## 5 Conclusions

Building on the access to two commercially available databases of default probability for firms, we develop a set of statistical tests to find out if those models indeed add value to the usual rating-grades published by rating agencies. We introduce the concept of Cumulative Accuracy Profile (CAP) that examines the percentage of companies defaulting after a year that were captured by the model. The CAP curve gives in one graph a full information on the quality of the prediction and we see that for the worst cases both models perform fairly well. Two other statistical studies: the timeliness of the prediction and the Miller's test confirm the CAP curves. In particular, Miller's test shows that both models add significant information to S&P rating, while the converse is less true.

The results of our study clearly demonstrate that the expected default probabilities provided both by KMV Credit Monitor and Moody's RiskCalc contain information that is not captured by the traditional rating and that they would signal risk of default faster (on average few months faster).

Finally and to answer the question that was raised in the title of this article, we can say that overall the credit risk models we have studied here do statistically fulfil most of their promises.

## 6 Acknowledgements

Gianluca Oderda would like to express his gratitude to Alessio Balduini and Bijan Khandani from Moody's Risk Management Services, for giving the authors free access to the RiskCalc database, and for constructive criticism of the early versions of the paper.

## References

- [1] Merton R., 1974, On the pricing of corporate debt: the risk structure of interest rate, *The Journal of Finance* **29**, 449-470.
- [2] Altman E., 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance* **23**, 589-609.
- [3] Crouhy M., D. Galai and R. Mark, 2000, A comparative analysis of current credit risk models, *Journal of Banking and Finance*, **24**, 57-117.
- [4] Kealhofer S., 1995, Managing default risk in portfolios of derivatives, in *Derivative Credit Risk, Advances in Measurement and Management*, published by Risk Publications, London, 49-66.
- [5] Credit Suisse, 1997, CreditRisk+: a credit risk management framework, Credit Suisse Financial Products papers.
- [6] Kealhofer S. and M. Kurbat, 2001, The default prediction power of the Merton approach, relative to debt ratings and accounting variables, KMV document released May 29, 2001.
- [7] Miller R., 1998, A non-parametric test for credit rating refinements, Risk Magazine August 1998 issue, a modified version is also available on the web at <http://home.earthlink.net/millerrisk/Papers/Nonparametric.htm>
- [8] Honerkamp J., 1993, Stochastic dynamical systems : concepts, numerical methods, data analysis, John Wiley & Sons.
- [9] Sobehart J.R., Keenan S., Stein R., Benchmarking Quantitative Default Risk Models: A Validation Methodology, Moody's Investors Service document released in March 2000.

# A Appendix: Model Predicted vs. Actual Defaults: Result Tables

In this Appendix we report the detailed results of the test discussed in Section 4.1 above. We start by reporting the results of KMV Credit Monitor. The default probability classes can be found in Table 4.

Obs.	KMV Predicted Defaults						Actual Defaults						$\chi^2$
	C.1	C.2	C.3	C.4	C.5	C.6	C.1	C.2	C.3	C.4	C.5	C.6	
22	0	0	1	3	4	6	0	0	0	5	2	6	2.41
23	0	0	1	3	3	5	0	0	0	6	3	6	2.41
24	0	0	1	3	4	5	0	0	0	9	3	7	4.62
25	0	0	1	4	3	5	0	0	1	8	5	7	2.88
26	0	0	1	4	4	5	0	0	1	7	6	11	3.95
27	0	0	1	4	4	6	0	0	1	6	9	11	4.14
28	0	0	1	4	5	8	0	0	1	6	8	15	3.58
29	0	0	1	5	9	12	0	0	1	2	6	24	5.75
30	0	0	1	5	9	13	0	0	0	1	8	25	7.86
31	0	0	1	5	11	14	0	0	0	3	7	27	6.5
32	0	0	1	6	9	16	0	0	0	1	4	33	12.2
33	0	0	1	5	9	19	0	0	0	1	4	35	10.4
34	0	0	1	5	9	16	0	0	0	1	7	26	7.05
35	0	0	1	6	9	19	0	0	0	1	3	29	9.69
36	0	0	1	6	9	19	0	0	0	0	4	29	11.38
37	0	0	1	6	9	16	0	0	0	1	7	28	8.25
38	0	0	1	6	8	15	0	0	0	3	6	27	5.78
39	0	0	1	5	8	14	0	0	1	2	8	25	4.72
40	0	0	1	6	8	14	0	0	0	2	6	25	6.1
41	0	0	1	5	9	15	0	0	0	3	9	21	3.15
42	0	0	1	6	9	16	0	0	0	2	6	25	5.59
43	0	0	1	6	9	18	0	0	0	3	4	25	5.38
44	0	0	1	5	10	18	0	0	0	4	4	30	7.25
45	0	0	1	5	9	18	0	0	0	4	7	28	4.51
46	0	0	1	6	11	18	0	0	2	2	11	28	4.21
47	0	0	1	6	12	17	0	0	2	3	10	28	3.99
48	0	0	1	6	10	19	0	0	2	1	11	29	6.39

Table 4: Predicted vs. realized defaults for KMV Credit Monitor.

For Moody's RiskCalc we find instead in Table 5:

Obs.	KMV Predicted Defaults						Actual Defaults						$\chi^2$
	C.1	C.2	C.3	C.4	C.5	C.6	C.1	C.2	C.3	C.4	C.5	C.6	
22	0	0	1	1	3	2	0	0	2	2	5	3	0.96
23	0	0	1	1	3	2	0	0	2	1	7	4	1.91
24	0	0	1	1	3	2	0	0	2	4	7	5	3.59
25	0	0	1	1	3	2	0	0	3	5	6	6	5.54
26	0	0	1	1	3	2	0	0	2	7	7	7	7.98
27	0	0	1	1	3	2	0	0	2	9	9	5	8.73
28	0	0	1	2	3	2	0	0	3	7	11	7	8.39
29	0	0	1	2	6	3	0	0	1	6	16	9	8.19
30	0	0	1	3	7	5	0	0	0	3	11	18	9.06
31	0	0	1	3	9	6	0	0	1	4	10	20	6.94
32	0	0	1	3	10	8	0	0	0	2	8	26	11.61
33	0	0	1	2	10	8	0	0	0	2	9	27	11.97
34	0	0	1	3	10	7	0	0	0	4	8	21	8.45
35	0	0	1	3	11	8	0	0	0	1	7	24	11.24
36	0	0	1	3	11	9	0	0	0	0	10	22	10.93
37	0	0	1	3	11	9	0	0	0	1	11	23	9.15
38	0	0	1	3	10	9	0	0	1	2	12	20	5.04
39	0	0	1	3	9	8	0	0	1	3	12	20	5.62
40	0	0	1	3	9	7	0	0	1	2	15	14	4.25
41	0	0	1	3	9	6	0	0	1	3	18	10	3.93
42	0	0	1	2	8	5	0	0	2	2	19	10	5.76
43	0	0	1	2	8	5	0	0	2	3	18	9	4.26
44	0	0	1	3	8	7	0	0	3	2	14	19	7.48
45	0	0	1	3	8	8	0	0	1	6	10	22	8.02
46	0	0	1	3	8	8	0	0	0	6	15	22	11.34
47	0	0	1	3	9	8	1	0	1	2	18	22	10.94
48	0	0	1	3	9	7	0	1	0	2	21	20	12.17

Table 5: Predicted vs. realized defaults for Moody's RiskCalc.

Finally we report the complete results of the test for model promptness, discussed in Section 4.3 above.

Company Name	S&P D-Time	KMV D-Time	$\Delta^{KMV}$	Moody's Caa-Time	$\Delta^{Moody's}$
ACME METALS	34	28	6	28	6
KAISER GROUP INTL INC	40	5	35	34	6
ANACOMP INC	55	1	54	51	4
ARMSTRONG HOLDINGS INC	56	55	1	55	1
BEN FRANKLIN RETAIL STORES	4	1	3	1	3
BK ENTERTAINMENT	38	15	23	29	9
BRED TECHNOLOGIES INC	42	34	8	31	11
CARMIKE CINEMAS INC -CL A	53	51	2	52	1
COHO ENERGY INC	36	33	3	35	1
AXIOHM TRANSACTION SOLUTIONS	44	31	13	32	12
EAGLE FOOD CENTERS INC	48	42	6	44	4
EMPLOYEE SOLUTIONS INC	49	26	23	35	14
METAL MANAGEMENT INC	56	30	26	31	25
GENESIS HEALTH VENTURES INC	48	36	12	44	4
GENEVA STL CO -CL A	35	21	14	31	4
GEOTEK COMMUNICATIONS	34	20	14	22	12
GRAHAM FIELD HEALTH PDS	45	36	9	44	1
IMPERIAL SUGAR CO	57	48	9	48	9
INTEGRATED HEALTH SVCS INC	44	35	9	42	2
AMERICAN BANKNOTE CORP	39	12	27	30	9
KCS ENERGY INC	40	29	11	30	10
KENETECH CORP	3	1	2	1	2
L A GEAR INC	34	20	14	20	14
LTV CORP	57	50	7	55	2
MARINER POST-ACUTE NTWRK INC	43	31	12	36	7
LOEHMANN'S INC	38	17	21	14	24
METROCALL INC	60	12	48	43	17
NORTHWESTERN STL & WIRE	57	16	41	34	23
PAGING NETWORK INC	47	31	16	13	34
PHONETEL TECHNOLOGIES INC	34	2	32	26	8
PILLOWTEX CORP	56	41	15	37	19
PRANDIUM INC	59	58	1	4	55
SAFETY-KLEEN	50	48	2	49	1
SALANT CORP	34	17	17	23	11
SERVICE MERCHANDISE CO	36	32	4	30	6
CAREMATRIX CORPORATION	56	42	14	44	12
BRAZOS SPORTSWEAR INC	34	27	7	28	6
TOKHEIM CORP	53	43	10	44	9
TOWER AIR INC	47	8	39	30	17
TOWN & COUNTRY CORP -CL A	34	1	33	13	21
TRANS WORLD AIRLINES	58	43	15	9	49
TULTEX CORP	38	33	5	33	5
CHIQUITA BRANDS INTL	58	55	3	58	0
DECORA INDS INC	56	44	12	45	11
INACOM CORP	51	36	15	51	0
WORLDTEX INC	57	36	21	37	20
SEMI-TECH CORP -CL A	42	11	31	13	29
PHILIP SERVICES CORP	34	28	6	31	3
AMERICAN ECO CORP	50	33	17	50	0
LAIDLAW INC	50	48	2	49	1
ROYAL OAK MINES INC	35	26	9	25	10
ICG COMMUNICATIONS	56	50	6	32	24
DRYPERS CORP	55	4	51	1	54
TRISM INC	40	1	39	30	10
SAFETY COMPONENTS INTL INC	46	42	4	44	2
LIVENT	34	32	2	32	2
FPA MEDICAL MGMT INC	34	26	8	27	7
US OFFICE PRODUCTS CO	60	29	31	30	30
CHS ELECTRONICS INC	49	4	45	37	12
AT TRACK COMMUNICATIONS INC	59	28	31	30	29
CROWN VANTAGE INC	48	12	36	11	37
PHYSICIANS RESOURCE GRP INC	34	29	5	32	2
FORCENERGY	36	29	7	30	6
WASTE SYSTEMS INTL INC	58	5	53	13	45
WIRELESS ONE INC	36	10	26	12	24
UNISON HEALTHCARE CORP	20	12	8	13	7
DECISIONONE HOLDINGS CORP	41	29	12	29	12
PLANET HOLLYWOOD INTL INC	37	38	-1	35	2
PARACEL'SUS HEALTHCARE CORP	47	21	26	31	16
AMERICAN PAD & PAPER CO	45	27	18	22	23
DAILEY INTL INC -CL A	39	29	10	32	7
SEABULK INTERNATIONAL INC	41	29	12	35	6
KITTY HAWK INC	50	40	10	49	1
STAGE STORES INC	51	34	17	36	15
STYLING TECHNOLOGY CORP	46	33	13	44	2
HURRICANE HYDROCARBONS LTD	38	29	9	34	4
UNIFORET INC	49	45	4	31	18
EAGLE GEOPHYSICAL INC	40	32	8	39	1
AMF BOWLING INC	54	31	23	35	19
USN COMMUNICATIONS INC	35	28	7	30	5
VLASIC FOODS INTERNATIONAL	58	47	11	46	12
LOEWS CINEPLEX ENTMT CORP	59	50	9	55	4
MASTER GRAPHICS INC	51	33	18	45	6
AMERN ARCHITECTURAL PDS INC	51	40	11	38	13
NORTHPOINT COMMUNICATIONS GP	58	56	2	44	14
ETOYS INC	60	56	4	56	4
PNV INC	56	50	6	52	4

Table 6: Model promptness at signalling default events: the Table reports for each of the default events in the time window analyzed, the time lag between the model signal and the actual default event.